

WHAT IS CLAIMED IS:

1. A system, comprising:
a storage device configured to store a plurality of files; and
5 a file system configured to manage access to said storage device, wherein said file system is configured to:
compute a compressed size of at least a portion of a given file; and
store an indication of said compressed size in data storage corresponding
to said given file.
10
2. The system as recited in claim 1, wherein said data storage comprises a named stream, and wherein said file system is further configured to store a compression dictionary corresponding to said at least a portion of said given file in said named stream subsequent to said computing.
- 15 3. The system as recited in claim 1, wherein said data storage comprises a named stream, wherein said given file includes a plurality of ordered portions, wherein said file system is further configured to store a respective plurality of compression dictionaries in said named stream, and wherein each said respective compression dictionary corresponds
20 to one of said ordered portions.
4. The system as recited in claim 3, wherein said at least a portion of said given file corresponds to a given ordered portion of said given file, and wherein said file system is further configured to:
25 suspend computing a compressed size of said given ordered portion of said given file; and
subsequent to said suspending, resume computing said compressed size of said given ordered portion of said given file without recomputing a compressed

size of a lower-ordered portion than said given ordered portion of said given file.

5. The system as recited in claim 3, wherein a given respective compression
5 dictionary corresponding to a higher-ordered portion of said given file includes a given respective compression dictionary corresponding to a lower-ordered portion.

6. The system as recited in claim 3, wherein said file system is further configured to:
detect a write operation to a given ordered portion of said given file;
10 in response to detecting said write operation, invalidate the respective compression dictionaries corresponding to said given ordered portion and any higher-ordered portions than said given ordered portion of said given file; and
subsequent to said invalidating, recompute a respective compressed size of
15 only said given ordered portion and any higher-ordered portions than said given ordered portion of said given file.

7. The system as recited in claim 1, wherein said file system is further configured to:
store a respective compressed size of each of said plurality of files in a
20 corresponding one of a plurality of respective named streams;
compute a compressed size of a concatenated file resulting from appending a first file to a second file; and
determine a value of a file harmony metric from said compressed size of said
concatenated file and said stored respective compressed sizes of said first
25 file and said second file.

8. A method, comprising:
storing a plurality of files;

computing a compressed size of at least a portion of a given file; and
storing an indication of said compressed size in data storage corresponding to said
given file.

5 9. The method as recited in claim 8, wherein said data storage comprises a named
stream, and further comprising storing a compression dictionary corresponding to said at
least a portion of said given file in said named stream subsequent to said computing.

10 10. The method as recited in claim 8, wherein said data storage comprises a named
stream, and further comprising storing a respective plurality of compression dictionaries
in said named stream, wherein said given file includes a plurality of ordered portions, and
wherein each said respective compression dictionary corresponds to one of said ordered
portions.

15 11. The method as recited in claim 10, wherein said at least a portion of said given
file corresponds to a given ordered portion of said given file, and further comprising:
suspending computing a compressed size of said given ordered portion of said
given file; and
subsequent to said suspending, resuming computing said compressed size of said
20 given ordered portion of said given file without recomputing a compressed
size of a lower-ordered portion than said given ordered portion of said
given file.

25 12. The method as recited in claim 10, wherein a given respective compression
dictionary corresponding to a higher-ordered portion of said given file includes a given
respective compression dictionary corresponding to a lower-ordered portion.

13. The method as recited in claim 10, further comprising:
detecting a write operation to a given ordered portion of said given file;
in response to detecting said write operation, invalidating the respective
compression dictionaries corresponding to said given ordered portion and
any higher-ordered portions than said given ordered portion of said given
file; and
subsequent to said invalidating, recomputing a respective compressed size of only
said given ordered portion and any higher-ordered portions than said given
ordered portion of said given file.
14. The method as recited in claim 8, further comprising:
storing a respective compressed size of each of said plurality of files in a
corresponding one of a plurality of respective named streams;
appending a first file to a second file to yield a concatenated file;
subsequent to said appending, computing a compressed size of said concatenated
file; and
determining a value of a file harmony metric from said compressed size of said
concatenated file and said stored respective compressed sizes of said first
file and said second file.
15. A computer-accessible medium comprising program instructions, wherein the
program instructions are computer-executable to:
store a plurality of files;
compute a compressed size of at least a portion of a given file; and
store an indication of said compressed size in data storage corresponding to said
given file.

16. The computer-accessible medium as recited in claim 15, wherein said data storage comprises a named stream, and wherein the program instructions are further computer-executable to store a compression dictionary corresponding to said at least a portion of said given file in said named stream subsequent to said computing.

5

17. The computer-accessible medium as recited in claim 15, wherein said data storage comprises a named stream, wherein the program instructions are further computer-executable to store a respective plurality of compression dictionaries in said named stream, wherein said given file includes a plurality of ordered portions, and wherein each
10 said respective compression dictionary corresponds to one of said ordered portions.

18. The computer-accessible medium as recited in claim 17, wherein said at least a portion of said given file corresponds to a given ordered portion of said given file, and wherein the program instructions are further computer-executable to:

15 suspending computing a compressed size of said given ordered portion of said given file; and
 subsequent to said suspending, resuming computing said compressed size of said given ordered portion of said given file without recomputing a compressed size of a lower-ordered portion than said given ordered portion of said
20 given file.

19. The computer-accessible medium as recited in claim 17, wherein a given
respective compression dictionary corresponding to a higher-ordered portion of said given file includes a given respective compression dictionary corresponding to a lower-ordered
25 portion.

20. The computer-accessible medium as recited in claim 17, wherein the program instructions are further computer-executable to:

detect a write operation to a given ordered portion of said given file;

in response to detecting said write operation, invalidate the respective

5 compression dictionaries corresponding to said given ordered portion and
any higher-ordered portions than said given ordered portion of said given
file; and

subsequent to said invalidating, recompute a respective compressed size of only

said given ordered portion and any higher-ordered portions than said given

10 ordered portion of said given file.

21. The computer-accessible medium as recited in claim 15, wherein the program instructions are further computer-executable to:

store a respective compressed size of each of said plurality of files in a

15 corresponding one of a plurality of respective named streams;

append a first file to a second file to yield a concatenated file;

subsequent to said appending, compute a compressed size of said concatenated
file; and

determine a value of a file harmony metric from said compressed size of said

20 concatenated file and said stored respective compressed sizes of said first
file and said second file.

25